

**ANALISIS SEGMENTASI PENJUALAN OBAT MENGGUNAKAN K-
MEANS DAN VISUALISASI STREAMLIT UNTUK OPTIMALISASI
PENGADAAN STOK PADA RUMAH SAKIT YPK MANDIRI**



**MARSA NABILA
NIM. 2110512048**

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA
2025**

**ANALISIS SEGMENTASI PENJUALAN OBAT MENGGUNAKAN K-
MEANS DAN VISUALISASI STREAMLIT UNTUK OPTIMALISASI
PENGADAAN STOK PADA RUMAH SAKIT YPK MANDIRI**

**MARSA NABILA
NIM. 2110512048**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA
2025**

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Marsa Nabila
NIM : 2110512048
Tanggal : 07 Juli 2025

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 07 Juli 2025

Yang Menyatakan,



Marsa Nabila

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Marsa Nabila
NIM : 2110512048
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SEGMENTASI PENJUALAN OBAT MENGGUNAKAN K-MEANS DAN
VISUALISASI STREAMLIT UNTUK OPTIMALISASI PENGADAAN STOK PADA
RUMAH SAKIT YPK MANDIRI**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta

Pada Tanggal : 07 Juli 2025

Yang Menyatakan,



Marsa Nabila

LEMBAR PENGESAHAN

Judul : Analisis Segmentasi Penjualan Obat Menggunakan K-Means Dan Visualisasi Streamlit Untuk Optimalisasi Pengadaan Stok Pada Rumah Sakit YPK Mandiri
Nama : Marsa Nabila
NIM : 2110512048
Program Studi : SI Sistem Informasi

Disetujui oleh :

Penguji 1:
Kraugusteecliana, S.Kom., M.Kom., MM.

Penguji 2:
Neny Rosmawarni, M.Kom.

Pembimbing 1:
Nur Hafifah Matondang, S.Kom, MM

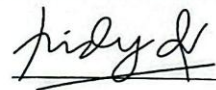
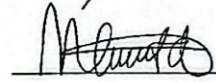
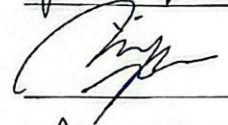
Pembimbing 2:
Nindy Irzavika, S.SI., M.T.

Diketahui oleh :

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTI.
NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM
NIP. 197605082003121002

Tanggal Ujian Tugas Akhir :
30 Juni 2025



ANALISIS SEGMENTASI PENJUALAN OBAT MENGGUNAKAN K-MEANS DAN VISUALISASI STREAMLIT UNTUK OPTIMALISASI PENGADAAN STOK PADA RUMAH SAKIT YPK MANDIRI

MARSA NABILA

ABSTRAK

Rumah Sakit YPK Mandiri memiliki prosedur pengadaan obat di rumah sakit telah terstruktur, proses pengelolaan persediaan masih berdasarkan data historis penjualan yang hanya berfokus pada pengecekan kuantitas tanpa mempertimbangkan pola penjualan tiap obat. Hal ini menyebabkan kurangnya prioritas yang tepat dalam perencanaan pengadaan, terutama ketika menghadapi lonjakan permintaan secara tiba-tiba. Pada penelitian ini diterapkan metode K-Means *Clustering* untuk melakukan segmentasi tingkat penjualan obat berdasarkan sejumlah fitur, yaitu nama obat, jumlah penjualan, bulan transaksi, koefisien variasi, dan jumlah bulan kemunculan obat. Proses segmentasi tersebut dilanjutkan dengan analisis mendalam terkait hubungan antara pola curah hujan dengan segmentasi penjualan, serta rekomendasi strategi pemasok guna mendukung pengadaan stok obat secara lebih efektif. Data yang digunakan merupakan data transaksi pada bulan Januari hingga Desember 2024 yang bersumber dari Rumah Sakit YPK Mandiri dan data curah hujan di Jakarta Pusat tahun 2024 yang bersumber dari data *online* Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Hasil pengujian, diperoleh jumlah klaster optimal adalah 3 dengan nilai *Silhouette Score* sebesar 0.4653, dan *Davies Bouldin Index* sebesar 0.7171 yang menunjukkan bahwa hasil *clustering* yang telah dilakukan dapat dikategorikan baik. Visualisasi data dengan Streamlit agar visualisasi menjadi lebih interaktif dan responsif yang dapat membantu pengguna dalam mengeksplor serta memahami hasil analisis lebih mendalam. Hasil dari penelitian ini dapat membantu dan mempermudah dengan menampilkan hasil visualisasi data melalui website dengan penerapan framework Streamlit, sehingga pihak rumah sakit dapat mengambil keputusan dan strategi pengadaan stok obat.

Kata Kunci : Rumah Sakit, Obat, K-Means, *Silhouette Score*, *Davies Bouldin Index*

ANALYSIS OF DRUG SALES SEGMENTATION USING K-MEANS AND STREAMLIT VISUALIZATION FOR OPTIMIZATION OF STOCK PROCUREMENT AT YPK MANDIRI HOSPITAL

MARSA NABILA

ABSTRACT

YPK Mandiri Hospital has a structured medicine procurement procedure in the hospital, the inventory management process is still based on historical sales data which only focuses on checking quantity without considering sales patterns of each medicine. This leads to a lack of proper prioritization in procurement planning, especially when facing sudden spikes in demand. In this study, the K-Means Clustering method is applied to segment medicine sales levels based on a number of features, namely medicine name, sales amount, transaction month, coefficient of variation, and number of months of medicine occurrence. The segmentation process is followed by an in-depth analysis of the relationship between rainfall patterns and sales segmentation, as well as recommendations for supplier strategies to support more effective medicine stock procurement. The data used is transaction data from January to December 2024 sourced from YPK Mandiri Hospital and rainfall data in Central Jakarta in 2024 sourced from online data from the Meteorology, Climatology and Geophysics Agency (BMKG). The test results obtained the optimal number of clusters is 3 with a Silhouette Score value of 0.4653, and a Davies Bouldin Index of 0.7171 which shows that the clustering results that have been carried out can be categorized as good. Data visualization with Streamlit to make the visualization more interactive and responsive which can help users explore and understand the results of the analysis more deeply. The results of this study can help and simplify by displaying the results of data visualization through the website with the application of the Streamlit framework, so that hospitals can make decisions and strategies for medicine stock procurement.

Keywords: *Hospital, Medicine, K-Means, Silhouette Score, Davies Bouldin Index*

KATA PENGANTAR

Segala puji dan rasa syukur penulis haturkan kepada Allah SWT karena berkat rahmat dan izin-Nya, penulis dapat menyelesaikan penulisan skripsi ini yang berjudul “Analisis Segmentasi Penjualan Obat Menggunakan Metode K-Means dan Visualisasi Streamlit Untuk Optimalisasi Pengadaan Stok Pada Rumah Sakit YPK Mandiri”.

Penulis menyadari bahwa dalam penyusunan skripsi ini, penulis banyak memperoleh bantuan, doa, serta dukungan dari banyak pihak. Oleh karena itu, pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua yang senantiasa mendoakan, memberikan dukungan, serta semangat dalam setiap langkah penulis, sehingga penulis dapat menyelesaikan studi hingga meraih gelar sarjana. Terima kasih atas doa-doa yang sangat luar biasa, kasih sayang, nasihat, dan motivasi yang selalu menguatkan penulis. Semoga Mamah dan Papah senantiasa diberi kesehatan dan kebahagiaan.
2. Kakek dan Nenek penulis, Alm. Bapak Usup & Ibu Nilah yang telah merawat dan membesarkan penulis dari kecil hingga dewasa, memberikan kasih sayang, perhatian, dan doa, sehingga penulis dapat terus berjuang dalam meraih mimpi dan cita-cita.
3. Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku Dekan Fakultas Ilmu Komputer UPNVJ.
4. Ibu Anita Muliawati, S.Kom., MTI., selaku Ketua Jurusan S1 Sistem Informasi Program Sarjana.
5. Ibu Nur Hafifah Matondang, S.Kom., M.M., M.T.I., selaku Dosen Pembimbing 1.
6. Ibu Nindy Irzavika, S.SI., M.T., selaku Dosen Pembimbing 2.
7. Ibu Neny Rosmawarni, M.Kom., selaku Dosen Pembantu yang telah memberikan penjelasan tambahan dan saran dalam penyempurnaan isi skripsi ini.
8. Pihak Rumah Sakit Umum YPK Mandiri karena telah bersedia menjadi objek untuk penelitian ini.
9. Sahabat dan teman terkasih penulis, Rahmi, Arsa, Wike, Tyas, Jihan, Iqlima, Hafshah, Fia, Ambar, Nafa, Almaas, Keysha, Elga, dan Manda yang telah setia memberikan dukungan, semangat, kerja sama, dan doa. Terima kasih telah menjadi teman bertumbuh dalam segala kondisi, menjadi pendengar yang baik, serta senantiasa memberikan semangat dan keyakinan kepada penulis bahwa setiap rintangan selama proses penyusunan skripsi ini dapat dilewati.

10. *Last but not least*, terima kasih untuk diri sendiri, karena telah bertahan dan berjuang sejauh ini. Terima kasih telah memilih untuk terus melangkah, meski tak jarang ingin menyerah. Terima kasih telah bertahan dalam malam-malam penuh tekanan, dalam rasa lelah yang tak terlihat, dalam tangis yang hanya diketahui oleh hati sendiri. Terima kasih telah berani menghadapi kegagalan dan percaya bahwa meski jalannya tidak selalu terang, satu langkah kecil setiap hari cukup untuk sampai ke tujuan. Perjalanan ini bukan hanya tentang menyelesaikan skripsi, tapi juga tentang tumbuh, belajar mengenal diri sendiri, menerima kelemahan, dan tetap berjalan dengan segala keterbatasan.

11. Seluruh pihak yang telah berkontribusi dalam membantu penulis yang tidak bisa disebutkan satu per satu, terimakasih atas doa dan dukungan yang sangat berharga.

Penulis telah melakukan segala upaya untuk menyelesaikan skripsi ini dengan sebaik mungkin, namun penulis menyadari bahwa masih ada berbagai kekurangan, baik dalam aspek isi maupun penulisan. Oleh karena itu, penulis sangat mengharapkan adanya saran dan kritik konstruktif akan sangat berharga.

Jakarta, 07 Juli 2025



Marsa Nabila

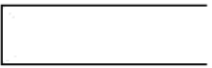
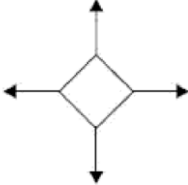
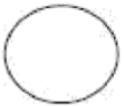
DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI.....	iii
DAFTAR SIMBOL	v
DAFTAR GAMBAR.....	vi
DAFTAR TABEL.....	vii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Ruang Lingkup Penelitian.....	4
1.4 Tujuan Penelitian	5
1.6 Luaran yang Diharapkan	5
1.7 Sistematika Penulisan.....	5
BAB 2. TINJAUAN PUSTAKA	7
2.1. Obat.....	7
2.2. Segmentasi Penjualan.....	7
2.3. Curah Hujan	8
2.4. <i>Data Mining</i>	9
2.5. Klasterisasi	12
2.6. Metode Elbow	14
2.7. K-Means	15
2.8. Evaluasi Model.....	16
2.9. Koefisien Variasi.....	18
2.10. Python	18
2.11. Streamlit	19
2.12. Penelitian Terdahulu	19
BAB 3. METODE PENELITIAN	32
3.1 Alur Penelitian	32
3.2 Tahapan Penelitian	32
3.2.1. Identifikasi Masalah	32
3.2.2. Persiapan Data	33
3.2.3. Pre-Processing Data.....	33
3.3 Implementasi Algoritma K-Means.....	35
3.4 Analisis Evaluasi Model	35
3.5 Visualisasi Data.....	36
3.6 Alat Bantu Penelitian	36
3.7 Jadwal Rencana Penelitian.....	37

BAB 4. HASIL DAN PEMBAHASAN	39
4.1. Persiapan Data.....	39
4.2. <i>Preprocessing</i> Data	40
4.2.1. <i>Data Integration</i>	40
4.2.2. <i>Data Selection</i>	41
4.2.3. <i>Data Cleaning</i>	42
4.2.4. <i>Data Transformation</i>	43
4.3. Implementasi K-Means <i>Clustering</i>	46
4.4. Evaluasi Model.....	52
4.5. Analisis dan Visualisasi Data	53
4.5.1. Analisis <i>Cluster</i>	53
4.5.2. Visualisasi Data	65
BAB 5. PENUTUP	71
5.1. Kesimpulan	71
5.2. Saran.....	72
DAFTAR PUSTAKA	73
DAFTAR RIWAYAT HIDUP	76
LAMPIRAN.....	77

DAFTAR SIMBOL

1. Simbol *Flowchart*

No.	Gambar	Nama	Keterangan
1		Terminal	Untuk menunjukkan awal dan akhir dari serangkaian proses.
2		Input/Ouput	Untuk menunjukkan cara kerja input/output.
3		Computer Processing	Untuk menunjukkan pemrosesan yang dilakukan oleh sistem komputer.
4		Predefined Processing	Untuk menunjukkan pemrosesan segala sesuatu yang tidak dinyatakan secara eksplisit dalam <i>flowchart</i> .
5		Comment	Untuk menangkap pernyataan penjelasan apa pun yang diperlukan untuk memperjelas sesuatu.
6		Flow Line	Untuk menghubungkan simbol-simbol.
7		Document	Digunakan saat input dan output dalam bentuk dokumen.
8		Decision	Untuk menunjukkan titik mana pun dalam proses pengambilan keputusan perlu ditentukan tindakan lebih lanjut.
9		On Page Connector	Untuk mengaitkan komponen-komponen flowchart yang dilanjutkan pada halaman yang sama.
10		Off Page Connector	Untuk mengaitkan komponen-komponen flowchart yang dilanjutkan pada halaman yang berbeda.

Sumber: (Chaudhuri, 2020)

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Data Mining (Raja dkk (2022)).....	11
Gambar 2. 2 Contoh <i>Clustering</i> (Jollyta et al., 2020).....	12
Gambar 2. 3 Tahapan K-Means <i>Clustering</i> (Jollyta et al., 2020).....	14
Gambar 2. 4 Hierarchical <i>Clustering</i> (Jollyta et al., 2020).....	14
Gambar 2. 5 Grafik Elbow Method (Jollyta et al., 2021)	15
Gambar 3. 1 Alur Penelitian.....	32
Gambar 4. 1 Dataset Setelah Data Integration.....	41
Gambar 4. 2 Dataset Setelah Data Selection	42
Gambar 4. 3 Data Duplikat	43
Gambar 4. 4 Dataset Setelah Data Cleaning	43
Gambar 4. 5 Dataset Setelah Dikelompokkan	44
Gambar 4. 6 Dataset Setelah Dinormalisasi	46
Gambar 4. 7 Metode Elbow Dalam Menentukan Jumlah <i>Cluster</i> Optimal.....	47
Gambar 4. 8 Rata-Rata Item Amount per <i>Cluster</i>	53
Gambar 4. 9 Rata-Rata Item Amount per <i>Cluster</i>	54
Gambar 4. 10 Rata-Rata CV per <i>Cluster</i>	55
Gambar 4. 11 Rata-Rata Jumlah Bulan Muncul per <i>Cluster</i>	56
Gambar 4. 12 <i>Top 10</i> Fungsi Obat (Use) pada <i>Cluster 1</i>	57
Gambar 4. 13 <i>Top 10</i> Fungsi Obat (Use) pada <i>Cluster 2</i>	58
Gambar 4. 14 <i>Top 10</i> Fungsi Obat (Use) pada <i>Cluster 3</i>	59
Gambar 4. 15 Distribusi Data per <i>Cluster</i>	60
Gambar 4. 16 Curah Hujan di Jakarta Pusat Tahun 2024.....	62
Gambar 4. 17 Rekapitulasi Total Penjualan Obat Berdasarkan Curah Hujan dan Bulan....	63
Gambar 4. 18 <i>Top 10</i> Obat Terlaris Berdasarkan <i>Cluster</i> , Curah Hujan, dan Bulan	64
Gambar 4. 19 Halaman Utama Dashboard	65
Gambar 4. 20 Hasil Klasterisasi dan Distribusi Persentase <i>Cluster</i>	66
Gambar 4. 21 Tampilan Data Per <i>Cluster</i>	66
Gambar 4. 22 Perbandingan Rata-rata Fitur per <i>Cluster</i>	67
Gambar 4. 23 <i>Top 10</i> Fungsi Obat per <i>Cluster</i>	67
Gambar 4. 24 Tampilan Halaman Rekomendasi	68
Gambar 4. 25 Visualisasi Data Curah Hujan.....	68
Gambar 4. 26 Visualisasi <i>Top 10</i> Item Obat per <i>Cluster</i> , Curah Hujan, dan Bulan	69
Gambar 4. 27 Visualisasi Rekapitulasi Total Penjualan Perbulan.....	70
Gambar 4. 28 Tampilan Ringkasan dan Kesimpulan	70

DAFTAR TABEL

Tabel 2. 1 Kategori Curah Hujan (BMKG, 2024)	9
Tabel 2. 2 Proses Algoritma K-Means (Id, 2021).....	16
Tabel 2. 3 Ringkasan Penelitian Terdahulu	19
Tabel 3. 1 Jadwal Rencana Penelitian.	37
Tabel 4. 1 Daftar Fitur pada Dataset Penjualan Obat	39
Tabel 4. 2 Daftar Fitur pada Dataset Curah Hujan	40
Tabel 4. 3 Nilai Sum of Squared Errors (SSE)	48
Tabel 4. 4 Centroid Awal.....	49
Tabel 4. 5 Distribusi Jumlah Data Setiap <i>Cluster</i>	51
Tabel 4. 6 Rata-rata Fitur per <i>Cluster</i>	51
Tabel 4. 7 Perbandingan Nilai SI dan DBI Antar Nilai K	52